

Reply to reviewer #1

We thank anonymous reviewer #1 for his/her constructive review that would improve the contents of our paper. The review comments by anonymous reviewer #1 are numbered and repeated below as *in italic letters*, followed by our answers. In the new draft with corrections (supplement file), red, green, purple, and blue corrections are the revisions suggested by reviewers #1, #2, #3, and co-authors, respectively.

<<Reviewer #1>>

<1-1> *The paper by Takeda et al reports the first attempt at a long term analysis of HFC-23 from ground-based FTIR. There are very compelling reasons to analyse this dataset based on the importance of HFC-23 as a strong greenhouse as (100 year GWP of over 12000), and the fact that it one of the HCFC family of chemicals whose production will not be phased out until 2030 under the Montreal Protocol for ozone protection. But its potential climate impact will remain for some time due to stockpiles and its long-lifetime. The addition of the NDACC network to long term measurements of this gas would be very welcome. However, this gas has not been a traditional target for this network as retrieving this gas from IR spectra is very difficult and challenging.*

We agree.

<1-2> *The purpose of this paper therefore is to outline, in some detail, how to retrieve total columns of HCFC-23 from IR spectra, obtain an annual trend and compare this trend against well-known and calibrated in-situ data. To retrieve this molecule is very difficult, so it will be of high interest to the NDACC FTIR community to follow this method. Indeed this will potentially unlock a worldwide dataset of measurements from Pole to Pole, linking in with the existing in-situ AGAGE network, and also offer a validation tool for satellite records.*

We agree.

<1-3> *The paper does indeed give a great deal of detail on the methodology, using one of the main software tools (SFIT4) used by the NDACC. This means the community can repeat this method without too many problems. However, the absorption signal from this species is so small (0.5%), and the spectra from filters 7/8 that are used throughout the network are relatively noisy, that without an independent dataset (AGAGE) the trends reported from this work would have been*

very difficult to justify.

We agree.

<I-4> The Japanese group leading this research have been involved in the NDACC for many years and therefore have significant experience in operating FTIR spectrometers and also analysing IR spectra.

Thank you.

<I-5> The manuscript itself is a mixed bag in terms of its use of English language. The first 6 or so pages of text, abstract, sections 1 and 2, flow well and have been written well. However after this, section 3 has many instances of poorly written English, and use of English terms that are inappropriately used. Given that there are a number of English speakers in the coauthorship list, it would seem that they have not carefully commented and/or made changes in the manuscript that are necessary. This is an important paper to get out into the community so this referee has gone to the trouble of pointing out most of the issues. Some may have been missed; so coauthors read carefully!

We are sorry for the poor English after section 3. The English-native co-authors of this paper checked the draft carefully until section 2, but they didn't check the draft so carefully after section 3. In this course of revision, our English-native co-author checked the revised draft carefully throughout to the end, and it is now reflected in the revised draft as blue-colored changes. Anyway, we thank Reviewer #1 for his/her great effort to check the English of this draft especially after section 3.

<I-6> This is an important paper for the FTIR community and a very useful outline of a method to analyse IR spectra for other researchers to follow. This paper therefore should be published, subject to the corrections/comments listed below

Thank you. Our corrections for your individual comments are shown below.

<I-7> Comments:

Section 1 and 2: good background and well written. One change; section 2.2, line 29 had => has.

It was corrected as suggested.

Section 3.1

1. Page 6, Line 20: "is a measurement noise." => "is the measurement noise."

It was corrected as suggested.

2. Line 23: "which are often so-called Jacobian" => "also called the Jacobians"

It was corrected as suggested.

3. Line 27: "is a gain matrix, whose line elements are so-called contribution function, which mean inversion sensitivity." => "is the gain matrix, or contribution function, which represents the sensitivity of the retrieved parameters to the measurement."

It was corrected as suggested.

4. Line 29: "relationship of" => "relationship"

It was corrected as suggested.

5. Page 7 line 2: "is a covariance" => "is the covariance"

It was corrected as suggested.

6. Line 6-7: "Comparing the equation (3), which is neglected the error terms of the forward model parameters and the measurement noise" => "Comparing equation (3), which neglects the error terms of the forward model parameters and the measurement noise"

It was corrected as suggested.

7. Line 9: "is usually non-linear problem," => "is usually a non-linear problem,"

It was corrected as suggested.

8. Line 16: "profile is found by the iteration" => "profile is found by iteration"

It was corrected as suggested.

9. Page 7 line 2: *“More detail is described in the following.” => “More detail is described in the following sections.”*

It was corrected as suggested.

10. Line 8: *“in maximum, about +5% relative “ => “about +5% relative”*

It was corrected as suggested.

11. Line 9: *“offset in measurement spectrum with second order polynomial” => “offset in the measured spectrum with a second order polynomial”*

It was corrected as suggested.

12. Line 11-12: *“On the other hand, the continuum level, which is equal to 100% in transmittance, was fitted by the following, because the shape of the continuum level is caused by the optical characterization of the FTIR instrument” => “On the other hand, the continuum level, which is equal to 100% in transmittance, has a shape that is caused by the optical characterization of the FTIR instrument”*

It was corrected as suggested.

13. Line 19: *“(the Fraunhofer lines),” => “(the Fraunhofer lines), and”*

It was corrected as suggested.

14. Line 26: *“has always been maintained best optical alignment.” => “has always been maintained with the best optical alignment.”*

It was corrected as suggested.

15. Page 9, line 26 : *“is temperature coefficient” => “is the temperature coefficient”*

It was corrected as suggested.

16. Line 29: *“in mid-infrared region” => “in the mid-infrared region”*

It was corrected as suggested.

17. Page 10, line 3: *“pressure data by National” => “pressure data obtained from the National”*

It was corrected as suggested.

18. Line 5: *“profiles by the COSPAR” => “profiles from the COSPAR”*

It was corrected as suggested.

19. Line 8: *“were basically used.” What does this mean exactly? Are there other profiles including HFC-23 that were not from WACCM?*

Individual species which used WACCM 40-years averaged profile is now shown here (N₂O, O₃, and PAN) and the word “basically” was removed. The paragraph describing CFC-12 and HCFCs a priori profile is moved before the paragraph of HFC-23 a priori description.

20. Line 9-10: *“For HFC-23, the WACCM does not compute its profile and thus a priori profile of HFC-23 was based on globally and annually mean mole fraction profile by two-dimensional”
=> “For HFC-23, WACCM does not provide a profile and thus the a priori profile of HFC-23 was based on the global and annual mean mole fraction profile by the two-dimensional”*

It was corrected as suggested.

21. Line 12: *“ppt at ground,” => “ppt at the ground,”*

It was corrected as suggested.

22. Line 14: *“ppt at ground” => “ppt at the ground”*

It was corrected as suggested.

23. Line 22: *“In retrieval of atmospheric” => “In the retrieval of an atmospheric”*

It was corrected as suggested.

24. Page 11, line 2: “respects” => “conserves”

It was corrected as suggested.

25. Lines 8-15. Please explain exactly how the Se “ad hoc” S/N was determined. In SFIT4 version 0.9.4.4 the actual S/N is recorded in the header of the input spectral file (commonly called the t15asc) for each micro-window. So while the Se matrix should in principle reflect the S/N for each spectral point, the implementation is such that a representative S/N for the whole window is applied. The ad hoc method is to scale this S/N to produce a retrieval that is both stable and has suitable dofs. This is part of the Steck procedure (4.D) to get α . What was the effective S/N for an α of 100?

Sorry we don’t have checked the SFIT4 core program in detail, but RMS value were used for Se in iteration process. We corrected the explanation for Se. S/N in t15asc is typically 2500 and effective S/N calculated from the RMS of the fit for α of 100 is typically 230-300.

26. Line 24: “which has been shown in the Atlas” => “as suggested in the NDACC IR reference microwindow Atlas”

It was corrected as suggested.

27. Line 26: “Since an H₂O absorption line having E'' of 586.48 cm⁻¹ is in the MW” => “Since a H₂O absorption line having an E'' of 586.48 cm⁻¹ is in the MW”

It was corrected as suggested.

28. Line 27: “temperature dependence of line strength are small” => “temperature dependence on the line strength are small”

It was corrected as suggested.

29. Page 12, line 2: “on the HITRAN” => “on HITRAN”

It was corrected as suggested.

30. Line 23: “*and fixed profile*” => “*and a fixed profile*”

It was corrected as suggested.

31. Page 13, line 3: “*values in profile*” => “*values in the profile*”

It was corrected as suggested.

32. Line 6: “*with the one*” => “*within one*”

It was corrected as suggested.

33. Line 7: “*with*” => “*within*”; both instances in this line.

It was corrected as suggested.

34. Line 18: “*for both all retrievals*” => “*for all retrievals*”

It was corrected as suggested.

35. Section 4.1: *on the question of the dofs for HCF-23. The regularization strength α is 100 which is quite strong. So this is a column scaling in effect, so the dofs will always end up being 1.0, regardless of how much information there is in the spectrum. There should be no surprise that the total column will be the only possible product from this method. That is, the dofs are being entirely driven by the method (Tikhonov), rather than the inherent spectral information.*

As discussed in Section 3.5, we used Tikhonov regularization rather than general optimal estimation method (OEM) in our sfit4 retrieval of HCF-23. Since we found that the DOFS was around 1.0 even if we apply OEM for the retrieval, we applied Tikhonov regularization method because it is more stable for the retrieval. The reason why we showed typical averaging kernel of HFC-23 retrieval is to show what altitudinal extent the HCF-23 retrieval has sensitivity.

36. Line 21: “*Subtract*” => “*Subtracting*”

It was corrected as suggested.

37. Line 28: *“from target gas” => “from the target gas”*

It was corrected as suggested.

38. Page 14, line 2: *“established well” => “well established”*

It was corrected as suggested.

39. Line 23-24: *“and the ones for other species were set to the values calculated from the used WACCM datasets.” => “and the uncertainties for other species were set to the values calculated from the appropriate WACCM datasets.”*

It was corrected as suggested.

40. Line 28: *“uncertainties on the” => “uncertainties reported from the”; NCEP has standard errors that are reported for their data which you can reference*

It was corrected as suggested.

41. Page 15 line 1: *“we set” => “we set to”*

It was corrected as suggested.

42. Line 4: *“and then their uncertainties” => “so that their uncertainties”*

It was corrected as suggested.

43. Line 5: *“more affects” => “has a larger effect on”*

It was corrected as suggested.

44. Line 6: *“error to” => “error on”*

It was corrected as suggested.

45. Line 12: *as mentioned ealier above, what is the ad hoc S/N exactly quantitatively?*

It was corrected as suggested.

46. Line 19: *“closed” => “close”*

It was corrected as suggested.

47. Lines 21&22: *“to the” => “on the”*

It was corrected as suggested.

48. Line 31: *the caption for table 4 is a little misleading. The word “uncertainty” is not correct, replace this with “error”*

It was corrected as suggested.

49. Page 16, line 17: *“observation (Vömel et al., 2007) which was executed” => “observations (Vömel et al., 2007) which were flown”*

It was corrected as suggested.

50. Line 26: *“and of a” => “and a”*

It was corrected as suggested.

51. Line 29: *“biases lead underestimation to the trend on the retrieved” => “biases lead to an underestimation of the trend on the retrieved”*

It was corrected as suggested.

52. Line 30: *“curvature is considered to the background correction.” => “curvature of the continuum is considered when applying the background correction.”*

It was corrected as suggested.

53. Page 17, line 17: “as the” => “as in the”

It was corrected as suggested.

54. Page 18, line 6: “to the HFC-23” => “to HFC-23”

It was corrected as suggested.

55. Page 19, line 27 “remaining of” => “remaining”

It was corrected as suggested.

56. Line 32: “region” => “regions”

It was corrected as suggested.

57. Page 20, line 4: “negative difference of average” => “negative average difference of”

It was corrected as suggested.

58. Line 10: “from 260 K to 300 K approximately.” => “from approximately 260 K to 300 K.”

It was corrected as suggested.

59. Line 11: “ppt as” => “ppt as the”

It was corrected as suggested.

60. Line 26: “of lower” => “of the lower”

It was corrected as suggested.

61. Line 27: “this should be done to further understand the reason for the negative bias and the seasonal cycle by intercomparison with the retrievals using the observed spectra at other” => “further study should be undertaken to understand the reasons for the negative bias and apparent

seasonal cycle by an intercomparison with HCF-23 total columns at other”

It was corrected as suggested.

62. Page 21, line 3: *“as green” => “as a green”*

It was corrected as suggested.

63. Lines 5, 7, 9,11: *“annual change rate” => “annual growth rate”*

It was corrected as suggested.

64. Line 14: *“Considering with the above, it is obviously indicated” => “Considering the above, it would seem”*

It was corrected as suggested.

65. Line 20: *“strategy basically agree well” this is a very vague assertion. How do they agree? Perhaps it is better to write something like “strategy agree well within the errors”*

It was corrected as suggested.

66. Line 21: *“These results indicated” => “These results indicate”*

It was corrected as suggested.

67. Page 22, line 22: *“from the Eurasia.” => “from Eurasia.”*

It was corrected as suggested.

68. Line 28: *“of lower” => “of the lower”*

It was corrected as suggested.

69. Lines 29,31: *“annual change rate” => “annual growth rate”*

It was corrected as suggested.

70. Line 33: “ground-based measurements.” => “ground-based in-situ measurements.”

It was corrected as suggested.

71. Page 31, table 4: It might be useful to add to this table the assumed uncertainties for each parameter that led to the computed error terms. Note the comment earlier about changing the caption. If however the uncertainties are included then the caption would read “errors and uncertainties”

We agree to the reviewer. Table 4 and its caption is modified to show assumed uncertainty for each parameter and random/systematic errors which are computed by the assumed uncertainty.

72. Page 38, figure 7: what exactly are the R values computed with, ie, the FTIR data against all in situ data for each site?

No. The R values in Figure 7 are correlation coefficients of linear fit of each dataset.

Reply to reviewer #2

We thank Dr. Chris Boone (reviewer #2) for his constructive review that would improve the contents of our paper. The review comments by reviewer #2 are numbered and repeated below as *in italic letters*, followed by our answers. In the new draft with corrections (supplement file), red, green, purple, and blue corrections are the revisions suggested by reviewers #1, #2, #3, and co-authors, respectively.

<<Reviewer #2 (Chris Boone)>>

<2-1> *Ground-based column HFC-23 observations are reported for two locations, retrieved from FTIR measurements using SFIT4. Error sources are evaluated, trends are determined and compared to results from the AGAGE network. The results from the Japanese station exhibit large seasonal fluctuations that are attributed to transport from Asia, but the Antarctic results show no significant seasonal variation. The wintertime (December-January-February) results from the Japanese station are 10-15% lower than the AGAGE 12-box model estimates, while the results from the Antarctic station are roughly 25% lower than AGAGE estimates. The discrepancies are attributed in part to deficiencies in the temperature dependence of the pseudolines employed in the calculation of HFC-23, coupled with the latitudinal differences for the Antarctic station. Trend values for the two stations show generally good agreement with results from equivalent AGAGE measurement sites, supporting the use of ground-based FTIR measurements for HFC-23 trend monitoring.*

Overall, the results appear to be of good quality, although I am not sure I agree with the interpretation of the seasonal HFC-23 variation (more on that later). The article is well organized, and the writing is clear, albeit with a few instances of wording/grammar issues.

Thank you.

<2-2> *I felt there was a missed opportunity here for placing the bias relative to the AGAGE results into a broader context. I agree with the authors' choice of using their winter (DJF) results as their "background level" to avoid the unexpected seasonal variation. On page 18, line 23, the authors state: values from December to February are mostly stable with a relatively small standard deviations of about +/-10% and a value of 10–15% smaller than the AGAGE in-situ measurements of HFC-23.*

Looking at Figure 4 of the Fernando et al article referenced in this paper, the global average results from the ACE-FTS exhibit very similar behavior, a persistent low bias relative to AGAGE

results, roughly 10% in magnitude. Because the ACE-FTS employs the Harrison HFC-23 cross sections in its retrievals and the current study uses pseudo lines based primarily on the same cross sections, this agreement in the bias relative to AGAGE does not seem like a coincidence. It looks like there is a systematic discrepancy between results derived using the Harrison cross sections and AGAGE results. I believe absolute calibration of the Harrison cross sections made use of the PNNL measurements, and I would be surprised if that group's determination of HFC-23 concentrations were off by 10% or so. However, I have no knowledge of the measurement technique employed by AGAGE, so I have no idea if a ~10% error on that end is feasible. It seems clear HFC-23 derived using the Harrison cross sections are inconsistent with AGAGE measurements, but I have no insight into the source of that discrepancy.

We agree to reviewer #2 on the systematic negative bias of HFC-23 derived using the Harrison cross sections (ACE-FTS) and pseudo lines (this study). Therefore, we added description of negative biases of -15 to -20% and -25%, and their possible cause due to the HFC-23 spectroscopy in the abstract.

<2-3> I am somewhat skeptical that the seasonal variation in measured HFC-23 is real. No other station sees such behavior. They theorize that it comes from long range transport, and yet not a hint of it continues on across the Pacific Ocean to Trinidad Head, despite the apparent large magnitude of the emissions. It is possible the station in question is located in just the right location downwind of a large emitter, and perhaps the flow would routinely divert away from Trinidad head as it travels across the Pacific (not my area of expertise). Emissions during HCFC-22 production is a major source of HFC-23, and Japan is a significant HCFC-22 producer, so perhaps the emissions originate from near the station, which would explain why they see such large quantities. However, I would be inclined to think there might be an artifact in the retrieval. Are there systematic differences in the residuals for baseline HFC-23 conditions versus enhanced HFC-23 conditions? If additional structure appears in the residuals in the latter case, the retrieval may be compensating for something missing in the analysis or compensating for large residuals from some other constituent (e.g., from water lines if H₂O levels are high). Is there a strong correlation of enhanced HFC-23 with H₂O, temperature, the ratio of HDO to H₂O, seasonal variations in CH₄, or biomass burning products like HCN or C₂H₆? I suspect H₂O might be the most dangerous of the bunch for impacting the retrieval. Correlation does not prove cause but may help identify the source of a retrieval artifact, if one exists.

Japanese HFC-23 emission had already diminished at around 2004 as shown in Figure 2-1. We think there is almost no emissions originate from near the station in Japan after 2004.

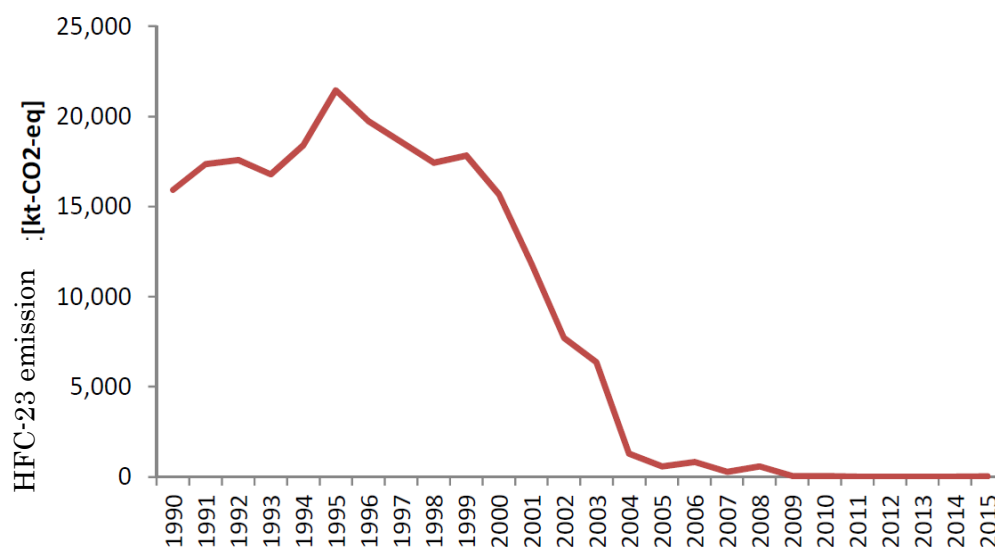


Figure 2-1 Temporal variation of HFC-23 emission in Japan from HCFC-22

[https://www.env.go.jp/earth/ondanka/ghg-mrv/methodology/material/methodology_2B9_1.pdf]

And you can find spikey enhancement of HFC-23 especially in summer season in the surface measurement at Ochiishi, Japan in Figure 2-2. (Ochiishi is located in the same island as Rikubetsu. You can get the figure from <https://gaw.kishou.go.jp/search/file/0053-2008-1502-01-01-9999>).

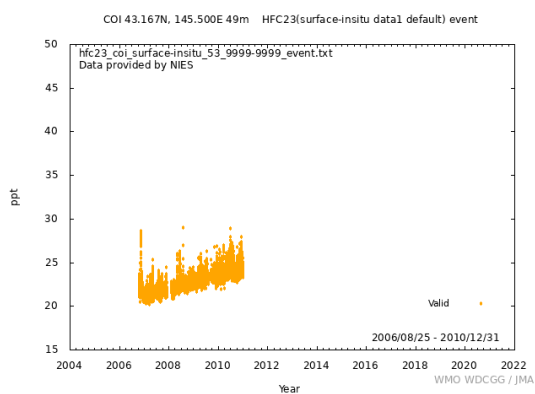


Figure 2-2. Surface HFC-23 measurement in Ochiishi

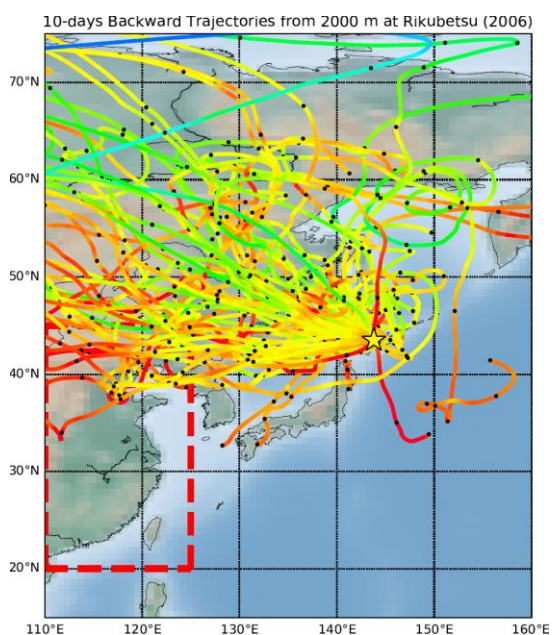


Figure 2-3. Airmass trajectory to Rikubetsu in 2006.

We calculated 10-days backward trajectories originated from Rikubetsu at 2000 km altitude on the days of FTIR observations. These backward trajectories showed that nearly 30 % of airmasses above Rikubetsu on the FTIR observation days came from Chinese region as is shown in Figure 2-3. We add these descriptions in Section 5.2. There is no significant difference in the RMS residuals between baseline HFC-23 conditions and enhanced HFC-23 conditions. Mean RMS residual and its standard deviation for baseline condition (DJF) in 2006-2010 are 0.296 ± 0.034 , while those for enhanced condition (other months) are 0.296 ± 0.048 . There is temperature dependency on the PLL and this explain 5% difference in maximum as described in Section 5.3. We have checked some correlations. Figures 2-4 and 2-5 are the correlations between HFC-23 and H₂O, and HFC-23 and HDO/H₂O, for Rikubetsu in 2006-2010, respectively. You can find positive correlations, but note that H₂O, HDO/H₂O, and also the transport from Eurasia are all high in summer season.

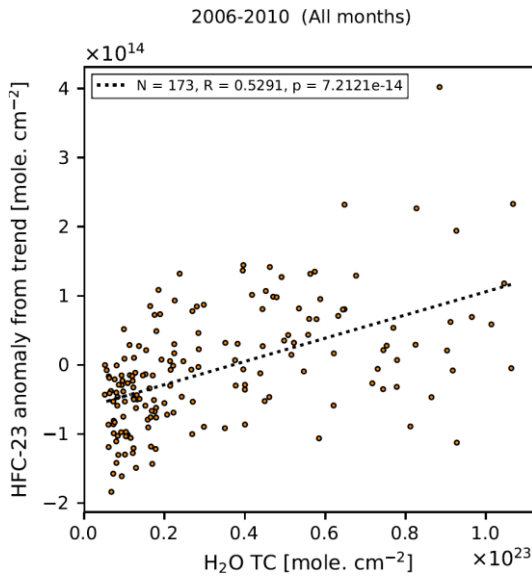


Figure 2-4. HFC-23 v.s. H₂O correlation

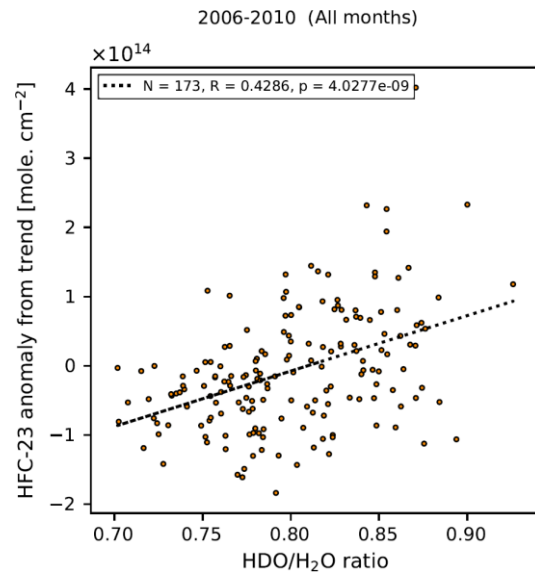


Figure 2-5. HFC-23 v.s. HDO/H₂O correlation

Figures 2-6 and 2-7 are the correlations between HFC-23 and H₂O, and HFC-23 and HDO/H₂O, for Syowa in 2007-2016, respectively. Here, there is no significant correlations. These results indicate that the retrieval artifacts from H₂O, and HDO/H₂O are not significant. We add description on this in Section 5.2.

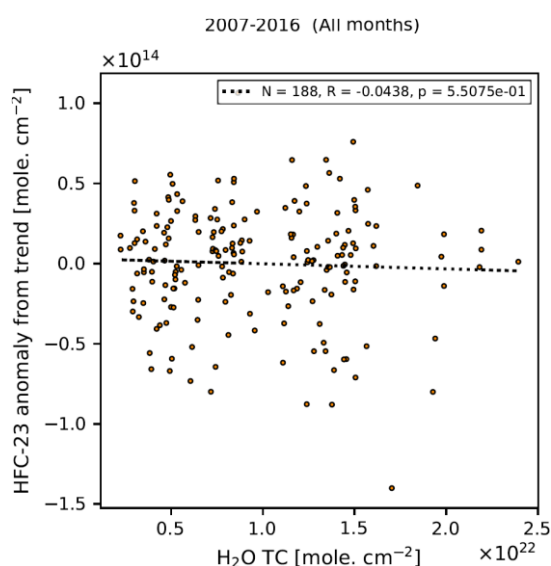


Figure 2-6. HFC-23 v.s. H₂O correlation.

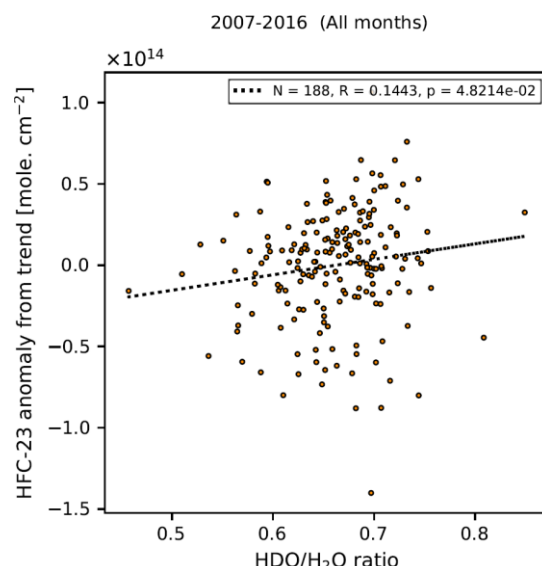


Figure 2-7. HFC-23 v.s. HDO/H₂O correlation

<2-4> *In my opinion, asserting that the HFC-23 enhancements were real would at minimum require comparing residuals for enhanced versus background conditions, ideally with a similar quality in both cases and no evidence of additional significant systematic features in the residuals for the enhanced case.*

As mentioned in the reply for the previous comment, there is no significant difference in the residuals between baseline HFC-23 conditions and enhanced HFC-23 conditions

<2-5> *The authors use two microwindows in the retrieval. I assume independent ‘background correction’ parameters (slope and curvature) are used in each window. Is it possible to have SFIT4 use a single set of background correction parameters spanning the two windows? The windows are close together, so I expect a unified curve for the background correction should be viable. This adjustment would presumably make the retrieval less susceptible to artifacts.*

Since we are using relatively wide two microwindows ranging 20 cm⁻¹ in total and 6 cm⁻¹ apart each other, we consider that using independent background correction parameters is better than using a single background correction parameter.

<2-6> *Page 8, line 25: for all observed spectra with the IFS-120/5HR instrument no ILS function was used because the instrument has always been maintained best optical alignment*

This statement is not entirely accurate. An instrumental line shape function is still required, associated with the FTS scan length and the finite field of view. It would be more appropriate

to say something like the ILS is accurately defined by the theoretical model for the given instrument configuration.

You are right. We modified the description as you suggested.

<2-7> Page 4, line 9: *However, the ACE-FTS observations do not have sensitivity to the troposphere where all HFC-23 emissions occur.*

Also not entirely accurate as stated. ACE-FTS measurements extend into the troposphere (i.e., below the tropopause). It would be more accurate to say ACE-FTS measurements do not extend low into the troposphere or do not probe near the surface.

You are right. Also reviewer #3 mentioned on this by his comment <3-15>. We finally deleted the description of tropospheric sensitivity issue from the draft.

<2-8> Page 20, line 3: *a discussion of temperature sensitivity in relation to the ACE-FTS.*

Note that ACE-FTS analysis uses the cross sections directly in the analysis (via a bilinear interpolation in pressure and temperature), whereas the current study employs pseudolines derived from a set of cross sections. I would assume that the deficiencies in temperature dependence described in the paper are a property of the pseudolines and not the cross sections themselves. There could conceivably be temperature consistency issues arising from the ACE-FTS interpolation approach, but I see no reason to expect it is similar in nature to the pseudoline temperature dependence deficiencies unless one claims the temperature dependence problems are inherent to the Harrison cross sections (I would need to see proof, if that were claimed).

You are right. We mixed up temperature dependence of pseudolines with that of cross sections. We deleted description of ACE-FTS issue here, and inserted it in the next paragraph.

<2-9> Page 7, line 25: *only about 1% of the atmospheric transmittance of solar infrared radiation at ground level (see Figure 3).*

It is traditional to make reference to figures in order. Figure 3 is referenced before Figures 1 and 2.

You are right. The reference for Figure 3 was deleted here.

<2-10> Page 17, line 21: *g is the column-averaged acceleration
acceleration due to gravity*

It was corrected as suggested.

<2-11> Grammar/wording issues:

>Page 3, line 6: *all UNFCCC CDM project were terminated ... [projects]*

It was corrected as suggested.

>Page 5, line 31: *detected with the MCT detector ... [measured with]*

It was corrected as suggested.

>Page 5, line 31: *Note, that ... [no comma]*

It was corrected as suggested.

>Page 11, line 31: *DOFS for HDO was increased when the wider window used ... [was used]*

It was corrected as suggested.

>Page 14, line 7: *due to the lack of vertically measurement data ... [vertically resolved]*

It was corrected as suggested.

>Page 14, line 31: *The SZA random uncertainty was assumed an uncertainty of 0.15° ... [assumed a value of]*

It was corrected as suggested.

>Page 15, line 18: *Since the MW for CH₄ pre-retrieval is closed to the HFC-23 MWs ... [close to]*

It was corrected as suggested.

>Page 16, line 29: *These relative biases lead underestimation to the trend ... [lead to an*

underestimation of the trend]

It was corrected as suggested.

>Page 19, line 32: in all temperature region ... [regions]

It was corrected as suggested.

<2-15> In conclusion, while I think it is a good paper, I would like to see more said on AGAGE results and results from studies employing the Harrison cross sections (this study and ACE-FTS work) exhibiting a fairly consistent bias, suggesting a systematic error somewhere. I would also like to see further proof that the apparent seasonal enhancement in HFC-23 is real (e.g., comparing residuals in the enhanced case versus the background case).

Thank you. We added some more descriptions on the negative bias issue, including the ACE-FTS negative bias, in the conclusion. As is described in the response to your comment <2-3>, we believe that the seasonal enhancement of HFC-23 is real. The description of backward trajectory calculation was also added in the conclusion.

Reply to reviewer #3

We thank anonymous reviewer #3 for his/her constructive review that would improve the contents of our paper. The review comments by anonymous reviewer #3 are numbered and repeated below as *in italic letters*, followed by our answers. In the new draft with corrections (supplement file), red, green, purple, and blue corrections are the revisions suggested by reviewers #1, #2, #3, and co-authors, respectively.

<<Reviewer #3>>

<3-1> *This paper describes the derivation of total column and dry-air mixing ratios of HFC-23, a compound of interest to the atmospheric science community because of its non-negligible contributions to radiative forcing, its long lifetime, and because controls exist on its emissions. Accurate retrievals of this chemical would be important for providing useful independent assessments of its atmospheric burden and how that has changed over time. It is clear that the challenge in providing accurate retrievals is a difficult one, given the very small absorbance that is involved, and because a number of other gases are potentially interfere. The authors have clearly considered many of the factors complicating this retrieval, but the manuscript could use some additional clarifications and considerations before publishing.*

Thank you. We try to make additional clarifications and considerations.

<3-2> *Suggesting that the apparent seasonality represents emissive influences is premature, in my opinion, without any discussions of sensitivities and likelihoods. Before this assertion seems at all possible, one would have to explore a number of things (inversion analysis not necessary):*

We try to add some additional discussion on the seasonality of the HFC-23 retrievals as follows.

<3-3> *Seasonal wind patterns reaching the site. Do they vary in a way that is potentially consistent with transport from potential source regions in Spring and not in the others? How about 2018-2019, where the seasonality seems much less pronounced. Given that emissions continue during these years, and perhaps predominantly from a region (China) close to Japan, one might expect larger seasonality in the more recent results, not reduced seasonality such as is observed.*

We calculated 10-days backward trajectories for all FTIR measurement days from Rikubetsu

originated at 2000 m. The attached Figures 3-1, 3-2, and 3-3 represents the 10-days backward trajectories for the FTIR measurement days in 2006, 2019, and 2020, respectively. We assumed that the rectangular area (20°N-45°N, 110°E-125°E) shown by red dotted lines in these figures represents the Chinese area, from where large HFC-23 emission is expected. As is shown in these figures, there are several trajectories which passed north part of Chinese area in 2006, while there is only one trajectory which passed Chinese area both in 2019 and 2020. The percentages of trajectories which passed Chinese area are: 29%, 5%, and 11%, respectively. This result explains why the analyzed HFC-23 don't show apparent seasonal variation in 2019 and 2020, when the FTIR observed only few air mass originated from China.

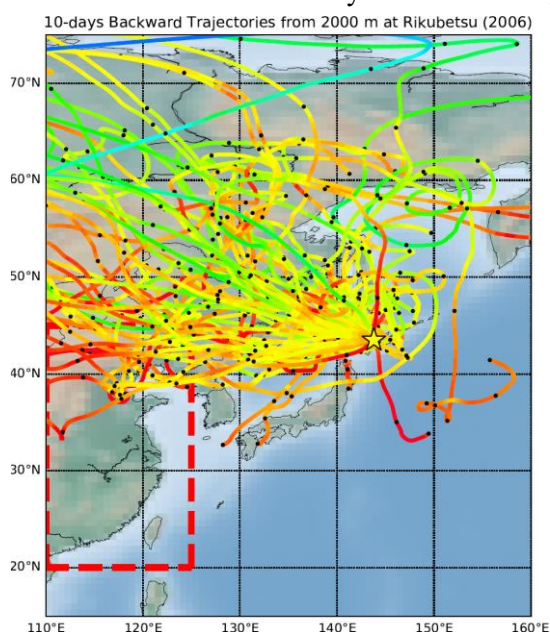


Figure 3-1. Trajectories in 2006

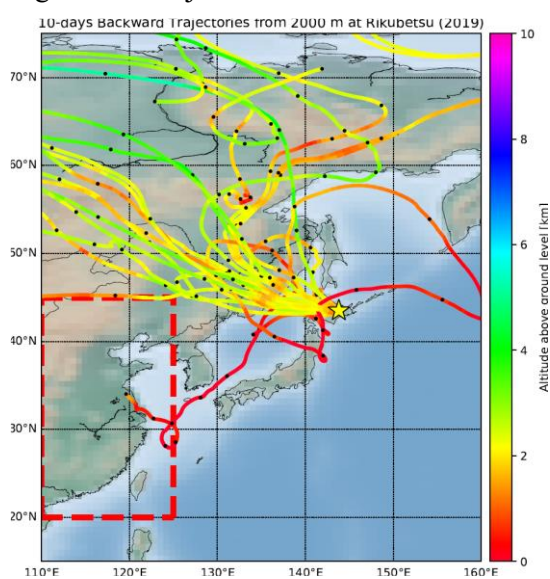


Figure 3-2. Trajectories in 2019

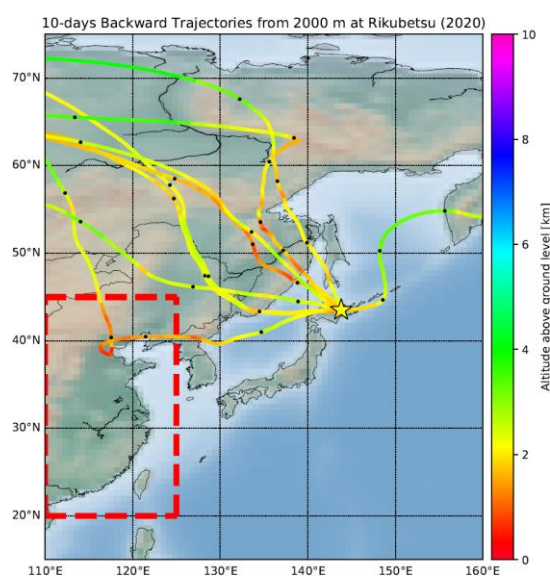


Figure 3-3. Trajectories in 2020

<3-4> While it is very good to see the discussion on a factor changing seasonally that affects the retrieved information (temperature), but its influence is insufficient to explain the unusually high amplitude that there is no hint of in the surface data. This is very puzzling. Why, for example, is the seasonality in the NH result much less in 2019-2020? Is it because the newer instrument is less susceptible to interfering influences, so in fact not the result of such a large emission?

As is explained in the reply to your comment <3-3> above, the backward 10-days trajectory in 2019 and 2020 did not often pass through Chinese area compared with 2006. We consider this is the major reason why we did not see large seasonality in 2019-2020.

<3-5> The seasonal changes in total column and mole fraction are on the order of a factor of 2. If this were truly an emissive signal, the consistency with which is observed during spring would enhanced mole fractions over a large region of the mid-latitude NH throughout an entire season and, hence, very large emissions. What emission magnitude would this demand, and it is reasonable given the global emissions derived for these years? Some qualitative discussion of these possibilities is warranted to determine if the hypothesis is, or is not, reasonable.

Although we are not quite familiar with the HFC-23 emission magnitude from Chinese area, we could see apparent differences among the spikes of surface HFC-23 measurements by AGAGE. The following three figures represent the continuous surface AGAGE measurement of HFC-23 at Fig. 3-4: Hateruma (HAT), Fig. 3-5: Cape Ochiishi (COI), and Fig. 3-6: Trinidad Head (THD). Since Hateruma station is quite close to China (about 500 km apart), it often captures the HFC-23-rich airmasses which is shown by the spiky dots in Fig. 3-4. Cape Ochiishi (43.2°N, 145.5°E), which is located rather close (about 150 km) to Rikubetsu where our FTIR is located, is about 2,000 km apart from China. Nevertheless, HFC-23 data in COI sometimes show spikes as shown in Fig. 3-5, which suggests medium-range transport of HFC-23 rich air-masses from China. The HFC-23 data in Trinidad Head (>10,000 km apart from China) seldom show spikes, which means that HFC-23 rich air was almost diluted after inter-continental transport across the Pacific.

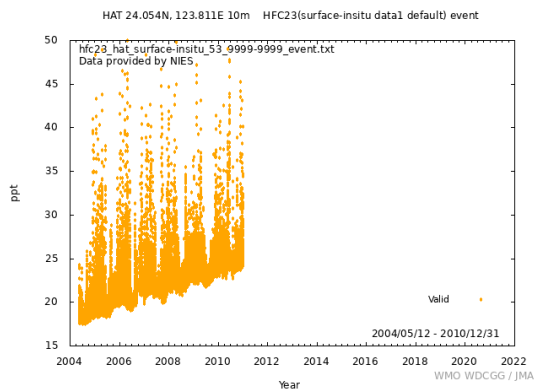


Fig. 3-4 HFC-23 measurement at Hateruma

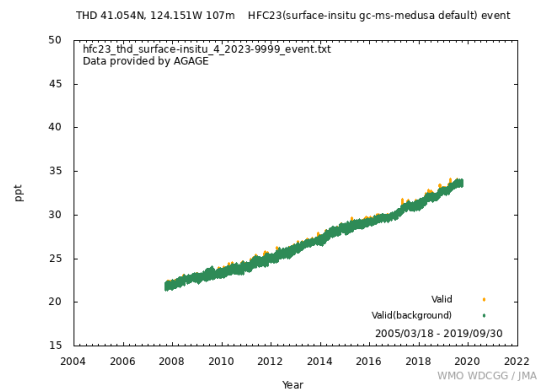


Fig. 3-6 HFC-23 measurement at Trinidad Head

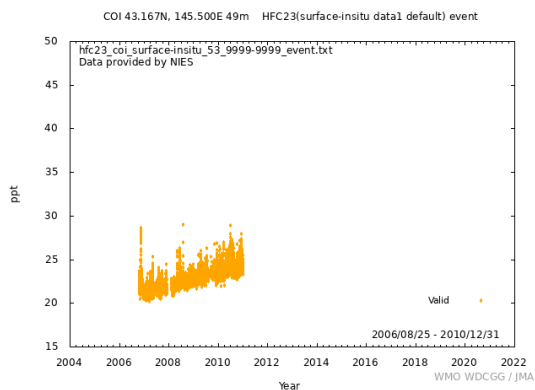


Fig. 3-5 HFC-23 measurement at Ochiishi

<3-6> On trends. It is not clear to the reader why it is important to express and assess trends for multiple periods and some further explanation on this point is needed, especially because it isn't apparent that there is a change in atmospheric growth rates corresponding to the chosen dates. Is it because of concurrent changes in instrumentation? Is it relevant to be deriving trends for significantly modified instruments? Especially over a span of time during which very few retrievals were made? The main conclusion is that this methodology provides accurate tracking of HFC-23 atmospheric mole fraction trends, yet there are significant differences in results and trends that aren't well caveated in this main conclusion.

The selection of trend period was not appropriate in the previous draft. We modified the selection of trend period more appropriately to represent the period when data were available. The Table 5 was modified to show three new periods. The contents in Section 5.4 was also modified. Now, the derived trends agreed within their standard deviation.

Other details:

<3-7> 19, line 14. This text is missing DJF, I believe: “indicated in Figure 8 that the retrievals at Rikubetsu have a negative bias”.

It is not appropriate to refer Figure 8 here, so it was deleted. The term ‘DJF’ was added as suggested.

<3-8> *An indication of the accuracy for the CH₄ pre-retrieval isn’t provided, but would be useful to understand to know if it is accurate or has biases that might affect the HFC-23 retrieval.*

The accuracy of the CH₄ pre-retrieval is written in the paragraph starting page 15, line 14. Since we are using relative absorption differences of CH₄ lines for the pre-retrieval and for the CH₄ line overlapped with HFC-23, their absolute accuracy or bias will not affect the HFC-23 retrieval.

<3-9> 12, lines 3-22 needs to more clearly written. Explain at first the different approaches that are available, and the terminology, so that the average reader will understand.

We modified the description of this paragraph to explain why we introduced ‘CH₄ pre-retrieval’ more clearly, at the beginning of this paragraph.

<3-10> *Consider subsetting data in some figures in an additional panel to allow the reader a better view of results that are relevant. Make clear, if true, that results in Figure 7 only include retained data with acceptable RMS values.*

We agree that old Figure 7 is too small and difficult to see individual points. We extended the vertical size of new Figure 7. Yes, the results in Figure 7 only include the retained data with acceptable RMS values as is describe in Section 5.1.1.

<3-11> *Abstract describes the work quite well, although more specificity and clarity is needed in lines 29-33. What is the size of the negative biases? Why are results from only some months used in the NH? On p.2, lines 4-5, having the capability to measure trends begs the long-term difference in NH results, [MORE HERE??]*

We added the description of amount of negative biases and their possible cause due to the HFC-23 spectroscopy in the abstract. The reason of the selection of DJF data at Rikubetsu is also

described. The advantage of addition of FTIR long-term measurement to fill the spatial and temporal gaps of the AGAGE observation is also added at the last of the abstract.

<3-12> 2, line 22. *Increasing emissions of HCFC-22 do not necessarily mean increasing emissions of HFC-23, given that HFC-23 is associated with the production of 22, not its emission, which is delayed by use in appliances. Focus on the tie between HCFC-22 production with HFC-23 emissions.*

We agree to the reviewer that HFC-23 increase is associated with the production of HCFC-22, not its emission. We deleted the word ‘Clearly’ here.

<3-13> 2, line 23-24. *ODPs have been calculated for HFCs, so they are in fact non-zero. The destruction arises from changes in the thermal structure of the atmosphere. Consider rephrasing “because they do not contain ozone depleting halogen atoms”.*

We agree. We modified the text as suggested.

<3-14> p.3, lines 12-13. *A citation is needed here. Are these truly mandated? Or are they aspirational goals to reduce emissions. Have India or China ratified the Kigali Amendment?*

There was a misunderstanding. India and China have not ratified the Kigali Amendment yet. We deleted the description of China and India from the draft.

<3-15> 4, lines 8-9. *This point does not seem all that relevant given the HFC-23 is a long-lived gas whose mole fraction is fairly evenly distributed.*

We agree to the reviewer. We deleted the description of tropospheric sensitivity issue from the draft.